Adjusted Expected Shortfall

Matteo Burzoni

Department of Mathematics, University of Milan, Italy

matteo.burzoni@unimi.it

Cosimo Munari

Center for Finance and Insurance and Swiss Finance Institute, University of Zurich, Switzerland

cosimo.munari@bf.uzh.ch

Ruodu Wang[1]

Department of Statistics and Actuarial Science, University of Waterloo, Canada

wang@uwaterloo.ca

July 14, 2020

Abstract

We introduce and study the main properties of a class of convex risk measures that refine Expected Shortfall by simultaneously controlling the expected losses associated with different portions of the tail distribution. The corresponding adjusted Expected Shortfalls quantify risk as the minimum amount of capital that has to be raised and injected into a financial position X to ensure that Expected Shortfall $\mathrm{ES}_p(X)$ does not exceed a pre-specified threshold $g(p)$ for every probability level $p \in [0,1]$. Through the choice of the benchmark risk profile g one can tailor the risk assessment to the specific application of interest. We devote special attention to the study of risk profiles defined by the Expected Shortfall of a benchmark random loss, in which case our risk measures are intimately linked to second-order stochastic dominance.

1 Introduction

In this paper we introduce and discuss the main properties of a new class of risk measures based on Expected Shortfall. Following the seminal paper by Artzner et al. (1999), we view a risk measure as a capital requirement rule. More precisely, we quantify risk as the minimal amount of capital that has to be raised and invested in a pre-specified financial instrument (which is typically taken to be risk free) to confine future losses within a pre-specified acceptable level of security. Value at Risk (VaR) and Expected Shortfall (ES) are the most prominent examples of risk measures in the above sense. If we fix a probability level $p \in [0,1]$ and model the future value of a financial position

[1]Ruodu Wang is supported by Natural Sciences and Engineering Research Council of Canada (RGPIN-2018-03823, RGPAS-2018-522590).

by a random variable X, then the VaR and ES of X at level p are respectively given by

$$\mathrm{VaR}_p(X) := \begin{cases} \inf\{x \in \mathbb{R} \mid \mathbb{P}(X \leqslant x) \geqslant p\} & \text{if } p \in (0,1], \\ \operatorname{ess\,inf} X & \text{if } p = 0, \end{cases}$$

$$\mathrm{ES}_p(X) := \begin{cases} \frac{1}{1-p}\int_p^1 \mathrm{VaR}_q(X)\mathrm{d}q & \text{if } p \in [0,1), \\ \operatorname{ess\,sup} X & \text{if } p = 1. \end{cases}$$

Here, we have adopted the convention to assign positive values to losses. In particular, note that $\mathrm{VaR}_p(X)$ is nothing but the (left) p-quantile of X. In line with our interpretation, the risk measures $\mathrm{VaR}_p(X)$ and $\mathrm{ES}_p(X)$ represent the minimal amount of cash that has to be raised and injected into X in order to ensure the following target solvency condition (for $0 < p < 1$):

$$\mathrm{VaR}_p(X) \leqslant 0 \iff \mathbb{P}(X \leqslant 0) \geqslant p,$$

$$\mathrm{ES}_p(X) \leqslant 0 \iff \int_p^1 \mathrm{VaR}_q(X)\mathrm{d}q \leqslant 0 \overset{F_X \text{ is continuous}}{\iff} \mathbb{E}[X|X \geqslant \mathrm{VaR}_p(X)] \leqslant 0,$$

where F_X is the cumulative distribution function of X. In words, the VaR solvency condition requires that the loss probability of X is capped by $1-p$ whereas the ES solvency condition that there is no loss on average beyond the (left) p-quantile of X. In the banking sector, the Basel Committee has recently decided to move from VaR at level 99% to ES at level 97.5% for the measurement of financial market risk. In the European insurance sector, VaR at level 99.5% is the reference risk measure in the Solvency II framework while ES at level 99% is the reference risk measure in the Swiss Solvency Test framework. In the past 20 years, an impressive body of research has investigated the relative merits and drawbacks of VaR and ES at both a theoretical and a practical level. This investigation led to a better understanding of the properties of the above two risk measures but also triggered many brand new research questions that go beyond VaR and ES themselves. We refer to early work on ES in Acerbi and Tasche (2002) Acerbi (2002), Frey and McNeil (2002), and Rockafellar and Uryasev (2002) (where ES was called Conditional VaR). Some recent contributions to the broad investigation on whether and to what extent VaR and ES meet regulatory objectives are Koch-Medina and Munari (2016), Embrechts et al. (2018), Weber (2018), Bignozzi et al. (2020), and Baes et al. (2020). For robustness problems concerning VaR and ES, see, e.g., Cont et al. (2010) and Krätschmer et al. (2014), and for their backtesting, see, e.g., Ziegel (2016), Du and Escanciano (2017), and Kratz et al. (2018).

A fundamental difference between VaR and ES is that, by definition, VaR is completely blind to the behaviour of the loss tail beyond the reference quantile whereas ES depends on the whole tail beyond it. It is often argued that this difference, together with the convexity property, makes ES a superior risk measure compared to VaR. In fact, this is the main motivation that led the Basel Committee to shift from VaR to ES in their market risk framework; see BCBS (2012). In the spirit of Bignozzi et al. (2020) and Mao and Wang (2020), the aim of this paper is to enhance our understanding of *how* tail risk is captured by ES. More specifically, being essentially an average beyond a given quantile, ES can only provide an aggregate estimation of risk which, by its very definition, does not distinguish across different tail behaviors with the same mean. We show how to capture this dimension of tail risk by introducing a new class of convex risk measures that includes ES as a particular case.

To best describe this new class, we start from the following simple example. Consider two normally distributed random variables $X_i \sim N(\mu_i, \sigma_i^2)$, with $\mu_1 = 1$, $\mu_2 = 0$, $\sigma_1 = 0.125$, $\sigma_2 = 0.5$. For every probability level $p \in (0, 1)$, the ES of the above random variables is explicitly given by

$$\mathrm{ES}_p(X_i) = \mu_i + \sigma_i \frac{\phi(\Phi^{-1}(p))}{1-p},$$

where ϕ and Φ are, respectively, the probability density and the cumulative distribution function of a standard normal random variable. For $p = 99\%$ the ES of both random variables is approximately equal to 1.33. In Figure 1 we plot the two distribution functions. Despite they have the same ES, the two risks are quite different mainly because of their different variance: The potential losses of X_1 tend to accumulate around its mean whereas those of X_2 are more disperse and can be significantly higher (compare the tails in Figure 2). A closer look at the ES profile of both random variables, i.e., at the function $p \mapsto \mathrm{ES}_p(X_i)$, shows that the ES profile of X_1 is more stable than that of X_2 (see Figure 2). In order to distinguish the above two risks, we introduce a comparative criterion that looks at the tail in a different way compared to ES. More specifically, we want a risk measure that is sensitive to changes in (any pre-specified portion of) the ES profile

$$p \mapsto \mathrm{ES}_p(X)$$

of a random variable X. To this end, we "adjust" ES by considering the new risk measure

$$\mathrm{ES}^g(X) := \sup_{p\in[0,1]} \{\mathrm{ES}_p(X) - g(p)\}$$

where $g : [0, 1] \to (-\infty, \infty]$ is a given increasing function. The choice of g allows to take into account any desired portion of the ES profile of X. The above adjusted ES is a monetary risk measure in the sense of Artzner et al. (1999). Indeed, the quantity $\mathrm{ES}^g(X)$ can be interpreted as the minimal amount of cash that has to be raised and injected into X in order to ensure the following target

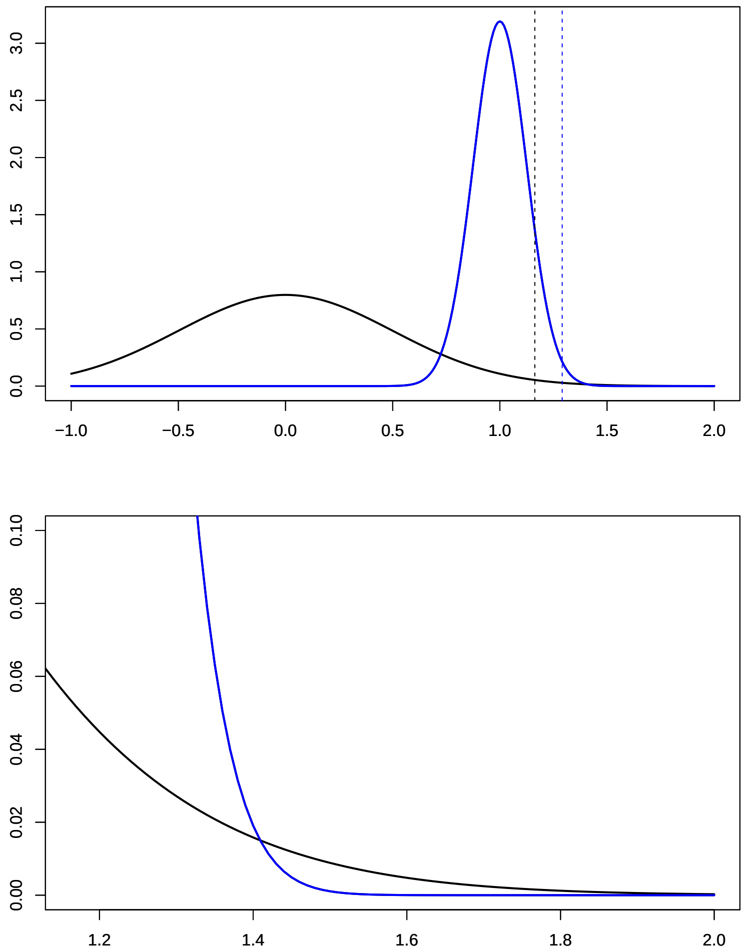


Figure 1: Above: Distribution of X_1 (blue) and X_2 (black). The vertical lines correspond to the respective 99% quantile. Below: Tails of of X_1 (blue) and X_2 (black) beyond the 99% quantile.

solvency condition:

$$\mathrm{ES}^g(X) \leqslant 0 \iff \mathrm{ES}_p(X) \leqslant g(p) \text{ for every } p \in [0,1].$$

To return to the above example, a simple way to distinguish X_1 and X_2 while, at the same time, focusing on average losses beyond the 99% quantile is to consider, e.g., the function

$$g(p) = \begin{cases} 0 & \text{if } p \in [0, 0.99], \\ 0.1 & \text{if } p \in (0.99, 0.9975], \\ \infty & \text{if } p \in (0.9975, 1]. \end{cases}$$

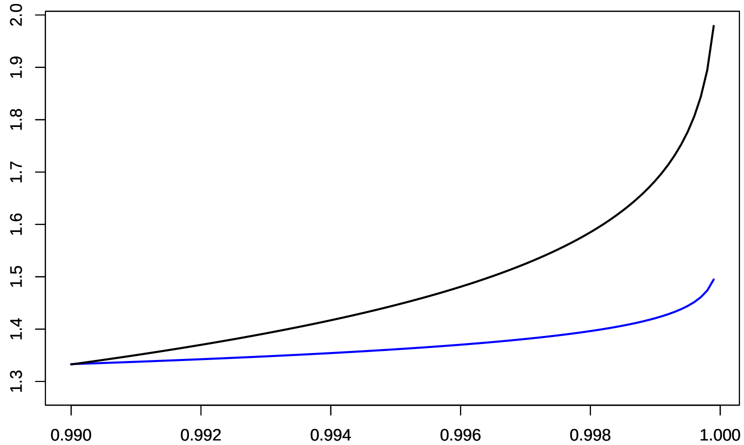


Figure 2: ES profile of X_1 (blue) and X_2 (black) for $p \geqslant 0.99$.

In this case, we easily obtain

$$\mathrm{ES}^g(X_i) = \max\{\mathrm{ES}_{0.99}(X_i), \mathrm{ES}_{0.9975}(X_i) - 0.1\} = \begin{cases} \mathrm{ES}_{0.99}(X_1) \approx 1.33 & \text{for } i = 1, \\ \mathrm{ES}_{0.9975}(X_2) - 0.1 \approx 1.45 & \text{for } i = 2. \end{cases} \tag{1}$$

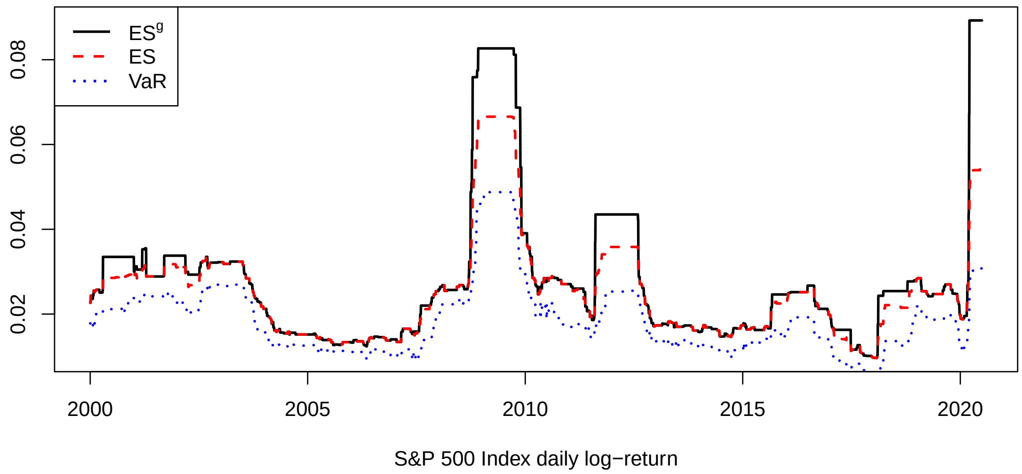
The focus of ES^g is still on the tail beyond the 99% quantile. However, the risk measure ES^g is able to detect the heavier tail of X_2 and penalize it with a higher capital requirement. This is because ES^g is additionally sensitive to the tail beyond the 99.75% quantile and penalizes any risk whose average loss on this far region of the tail is too large.

For the sake of illustration, we use a similar target risk profile to compare the behavior of the classical ES and the adjusted ES on real data. We collect the S&P 500 and the NASDAQ Composite indices daily log-returns (using closing prices) from January 01, 1999 to June 30, 2020. Each index has 5406 data points (publicly available from Yahoo Finance). We simply produce empirical values of the risk measures using a rolling window of one year without assuming any time-series models. More specifically, at each day (starting from Jan 2000), the risk measures are estimated from the empirical distribution of the negative log-return (log-loss) in the preceding year. We consider

$$g(p) = \begin{cases} 0 & \text{if } p \in [0, 0.95], \\ 0.01 & \text{if } p \in (0.95, 0.99], \\ \infty & \text{if } p \in (0.99, 1], \end{cases}$$

which yields

$$\mathrm{ES}^g(X) = \max\{\mathrm{ES}_{0.95}(X), \mathrm{ES}_{0.99}(X) - 0.01\}$$



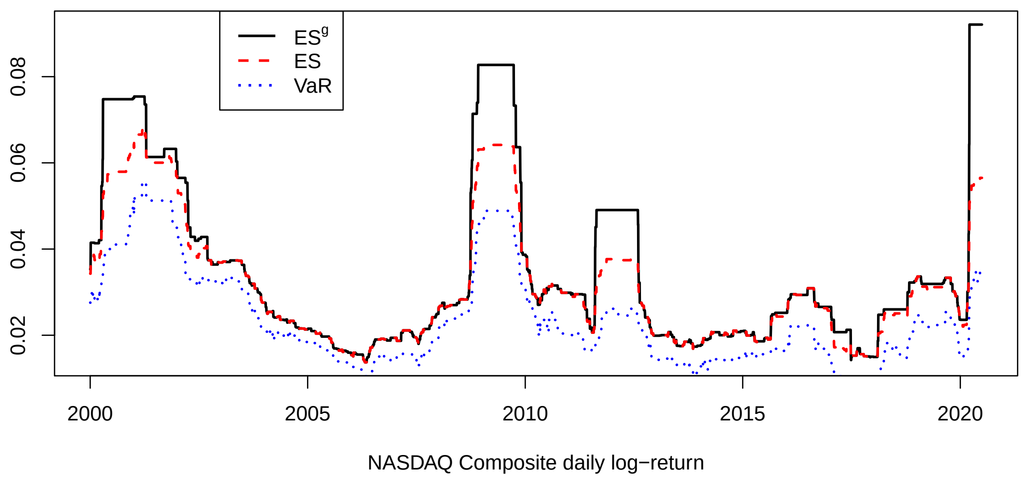


Figure 3: Empirical $\mathrm{ES}_{0.95}$, ES^g and $\mathrm{VaR}_{0.95}$ for S&P 500 and NASDAQ.

similar to (1) in a different context. The numbers 0.95, 0.99, and 0.01 that appear in g are chosen for the ease of illustration only. The 20-year empirical values of ES at level 95% and ES^g, as well as those of VaR at level 95%, are plotted in Figure 3. As we can see from the numerical results on both S&P 500 and NASDAQ, the estimated values of ES^g and the reference ES agree with each other during most of the considered time horizon. However, during periods of significant financial stress, such as the dot-com bubble in 2000, the subprime crisis in 2008, and the COVID-19 crisis in early 2020, ES^g is visibly larger than the reference ES. This illustrates that ES^g may capture tail risk in a more appropriate way than ES, especially under financial stress.

As illustrated above, a key feature of adjusted ES is the flexibility in the choice of the target risk profile g. Indeed, the same random loss can be considered more or less relevant depending on a variety of factors, including the availability of hedging strategies or other risk mitigation tools in the underlying business sector. The choice of g can be therefore tailored to the particular area of application by assigning different weights to different portions of the reference tail. We illustrate this feature by discussing a simple stylized example in the context of cyber risk. Differently from other operational risks, cyber risk has a strong geographical component. The empirical study Biener et al. (2015), which takes into account 22,075 incidents reported between March 1971 and September 2009, reveals that "Northern America has some of the lowest mean cyber risk and non-cyber risk losses, whereas Europe and Asia have much higher average losses despite Northern American companies experience more than twice as many (51.9 per cent) cyber risk incidents than European firms (23.2 per cent) and even more than twice as many as firms located on other continents". A possible reason is that North American companies may be better equipped to protect themselves against such events. Cyber risk cannot be properly managed by a simple frequency-severity analysis. In the qualitative analysis of Refsdal et al. (2015), many additional factors are identified including:

- Ease of discovery: How easy is it for a group of attackers to discover a given vulnerability?
- Ease of exploit: How easy is it for a group of attackers to actually exploit a given vulnerability?
- Awareness: How well known is a given vulnerability to a group of attackers?
- Intrusion detection: How likely is an exploit to be detected?

The answers may very well depend on the specific sector if not on the specific firms under consideration. The choice of different reference risk profiles g across companies might be a way to apply the theory of risk measures in the spirit of Artzner et al. (1999) to the rather complex analysis of this type of risk. For example, it would be possible to set

$$g(p) = \begin{cases} \mathrm{ES}_{0.99}(Z_1) & \text{if } p \in [0, 0.99], \\ \mathrm{ES}_{0.999}(Z_2) & \text{if } p \in (0.99, 0.999], \\ \mathrm{ES}_{0.9999}(Z_3) & \text{if } p \in (0.999, 0.9999], \\ \infty & \text{otherwise,} \end{cases}$$

where Z_1, Z_2, Z_3 are suitable benchmark random losses. The resulting adjusted ES is

$$\mathrm{ES}^g(X) = \max\{\mathrm{ES}_{0.99}(X) - \mathrm{ES}_{0.99}(Z_1), \mathrm{ES}_{0.999}(X) - \mathrm{ES}_{0.999}(Z_2), \mathrm{ES}_{0.999}(X) - \mathrm{ES}_{0.999}(Z_3)\}.$$

The associated target solvency condition is given by

$$\mathrm{ES}^g(X) \leqslant 0 \iff \begin{cases} \mathrm{ES}_{0.99}(X) \leqslant \mathrm{ES}_{0.99}(Z_1), \\ \mathrm{ES}_{0.999}(X) \leqslant \mathrm{ES}_{0.999}(Z_2), \\ \mathrm{ES}_{0.9999}(X) \leqslant \mathrm{ES}_{0.9999}(Z_3). \end{cases}$$

The choice of g should be motivated by specific cyber risk events (see Refsdal et al. (2015) for a categorization of likelihood/severity for different cyber attacks): The one in a hundred times event could be the malfunctioning of the server, the one in a thousand times event the stealing of the profile data of the clients, the one in a million times event the stealing of the credit cards details of the customers. Note that it is possible to choose a single benchmark random loss or a different benchmark random loss for each considered incident. This choice could also be company specific so as to reflect the company's ability to react to the different types of cyber attacks. This is in line with Biener et al. (2015), which says that "Regarding size (of the average loss per event), we observe a U-shaped relation, that is, smaller and larger firms have higher costs than medium-sized. Possibly, smaller firms are less aware of and less able to deal with cyber risk, while large firms may suffer from complexity".

The goal of this paper is to introduce the class of adjusted ES's and discuss their main theoretical properties. In Section 2 we provide a formal definition and a useful representation of adjusted ES's together with a broad discussion on some basic properties. A special interesting case is when the risk profile g is given by the ES of a benchmark random variable. We focus on this situation in Section 3 and show that such special adjusted ES's are strongly linked with second-order stochastic dominance. More precisely, they coincide with the monetary risk measures for which acceptability is defined in terms of carrying less risk, in the sense of second-order stochastic dominance, than a given benchmark random variable. Despite the importance of such a concept, we are not aware of earlier attempts to explicitly construct monetary risk measures based on second-order stochastic dominance. In Section 4 we focus on a variety of optimization problems featuring risk functionals either in the objective function or in the optimization domain and study the existence of optimal solutions in the presence of this type of risk measures. In each case of interest we are able to establish explicit optimal solutions.

2 Adjusting ES via target risk profiles

Throughout the paper we fix an atomless probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and denote by $\mathcal{X}$ the space of (equivalent classes with respect to $\mathbb{P}$-almost sure equality of) $\mathbb{P}$-integrable random variables. For any two random variables $X, Y \in \mathcal{X}$ we write $X \sim Y$ whenever X and Y are identically distributed. We adopt the convention that positive values of $X \in \mathcal{X}$ correspond to losses. Recall that VaR is defined as

$$\mathrm{VaR}_p(X) := \begin{cases} \inf\{x \in \mathbb{R} \mid \mathbb{P}(X \leqslant x) \geqslant p\} & \text{if } p \in (0,1], \\ \operatorname{ess\,inf} X & \text{if } p = 0, \end{cases}$$

and ES is defined as

$$\mathrm{ES}_p(X) := \begin{cases} \frac{1}{1-p}\int_p^1 \mathrm{VaR}_q(X)\mathrm{d}q & \text{if } p \in [0,1), \\ \operatorname{ess\,sup} X & \text{if } p = 1. \end{cases}$$

The focus of the paper is on the following class of risk measures. Here and in the sequel, we denote by $\mathcal{G}$ the set of all functions $g : [0,1] \to (-\infty, \infty]$ that are increasing (in the non-strict sense) and not identically ∞. Moreover, we use the convention $\infty - \infty = -\infty$.

Definition 2.1. Consider a function $g \in \mathcal{G}$ and define the set

$$\mathcal{A}_g := \{X \in \mathcal{X} \mid \forall p \in [0,1],\ \mathrm{ES}_p(X) \leqslant g(p)\}.$$

The functional $\mathrm{ES}^g : \mathcal{X} \to (-\infty, \infty]$ defined by

$$\mathrm{ES}^g(X) := \inf\{m \in \mathbb{R} \mid X - m \in \mathcal{A}_g\}.$$

is called the *g-adjusted Expected Shortfall (g-adjusted ES).*

To best appreciate the financial meaning of the above risk measures, it is useful to consider the *ES profile* associated with a random variable $X \in \mathcal{X}$, i.e., the function

$$p \mapsto \mathrm{ES}_p(X).$$

From this persepctive, the function g in the preceding definition can be interpreted as a threshold between acceptable (safe) and unacceptable (risky) ES profiles. In this sense, the set $\mathcal{A}_g$ consists of all the positions with acceptable ES profile and the quantity $\mathrm{ES}^g(X)$ represents the minimal amount of capital that has to be injected into X in order to align its ES profile with the chosen acceptability profile. For this reason, we will sometimes refer to g as the target *risk profile*.

The above risk measure is a natural extension of ES. Indeed, if for each $p \in [0,1]$ we define

$$g(q) = \begin{cases} 0 & \text{if } q \in [0,p], \\ \infty & \text{if } q \in (p,1], \end{cases}$$

we obtain that $\mathrm{ES}_p(X) = \mathrm{ES}^g(X)$ for every random variable $X \in \mathcal{X}$.

The next proposition highlights an equivalent but operationally preferable formulation of adjusted ES's which also justifies the chosen terminology.

Proposition 2.2. *For every risk profile $g \in \mathcal{G}$ and for every $X \in \mathcal{X}$ we have*

$$\mathrm{ES}^g(X) = \sup_{p\in[0,1]} \{\mathrm{ES}_p(X) - g(p)\}.$$

Proof. Fix $X \in \mathcal{X}$ and note that for every $m \in \mathbb{R}$ the condition $X - m \in \mathcal{A}_g$ is equivalent to

$$\mathrm{ES}_p(X) - m = \mathrm{ES}_p(X - m) \leqslant g(p)$$

for every $p \in [0,1]$. For $p = 1$ both sides could be equal to ∞. However, in view of our convention $\infty - \infty = -\infty$, the above inequality holds if and only if $m \geqslant \mathrm{ES}_p(X) - g(p)$ for every $p \in [0,1]$. The desired representation easily follows. □

Remark 2.3. (i) The above definition is reminiscent of the definition of *Loss Value at Risk* (LVaR) in Bignozzi et al. (2020). In that case, one takes an increasing and right-continuous function $\alpha : [0,\infty) \to [0,1]$ (the so-called benchmark loss distribution) and defines the acceptance set by

$$\mathcal{A}_\alpha := \{X \in \mathcal{X} \mid \mathbb{P}(X > x) \leqslant \alpha(x),\ \forall x \geqslant 0\}.$$

The corresponding LVaR is given by

$$\mathrm{LVaR}_\alpha(X) := \inf\{m \in \mathbb{R} \mid X - m \in \mathcal{A}_\alpha\}.$$

The quantity $\mathrm{LVaR}_\alpha(X)$ represents the minimal amount of capital that has to be injected into the position X in order to ensure that, for each loss level x, the probability of exceeding a loss of size x is controlled by $\alpha(x)$. According to Proposition 3.6 in the cited paper, we can equivalently write

$$\mathrm{LVaR}_\alpha(X) = \sup_{p\in[0,1]} \left\{\mathrm{VaR}_p(X) - \alpha_+^{-1}(p)\right\}, \tag{2}$$

where α_+^{-1} is the right inverse of α. This highlights the similarity with adjusted ES's.

(ii) It is clear, see also below, that ES^g is monotonic with respect to second-order stochastic dominance. This implies that ES^g belongs to the class of *consistent risk measures* as defined in Mao and Wang (2020). In fact, Theorem 3.1 in that paper shows that any consistent risk measure can be expressed as an infimum of ES^g's over a suitable class of risk profiles g's. In this sense, adjusted ES's can be seen as the building blocks for risk measures that are consistent with second-order stochastic dominance. This class is very large and includes all law-invariant convex risk measures.

The remainder of this section is devoted to discussing some basic properties of adjusted ES's. It follows immediately from our definition that every adjusted ES is a monetary risk measure in the sense of Föllmer and Schied (2016), i.e., is monotone and cash additive. The other properties listed below are automatically inherited from the corresponding properties of ES. For any random variables $X, Y \in \mathcal{X}$ we say that X *dominates* Y *with respect to second-order stochastic dominance*, written $X \geqslant_{\mathrm{SSD}} Y$, whenever $\mathbb{E}[u(-X)] \geqslant \mathbb{E}[u(-Y)]$ for every increasing and concave function $u : \mathbb{R} \to \mathbb{R}$.

In the language of utility theory, this means that X is preferred to Y by every risk-averse agent (recall that positive values of X and Y represent losses).

Proposition 2.4. *For every risk profile $g \in \mathcal{G}$ the risk measure ES^g satisfies the following properties:*

(1) monotonicity*: $\mathrm{ES}^g(X) \leqslant \mathrm{ES}^g(Y)$ for all $X, Y \in \mathcal{X}$ such that $X \leqslant Y$.*

(2) cash additivity*: $\mathrm{ES}^g(X+m) = \mathrm{ES}^g(X) + m$ for all $X \in \mathcal{X}$ and $m \in \mathbb{R}$.*

(3) law invariance*: $\mathrm{ES}^g(X) = \mathrm{ES}^g(Y)$ for all $X, Y \in \mathcal{X}$ such that $X \sim Y$.*

(4) convexity*: $\mathrm{ES}^g(\lambda X + (1-\lambda)Y) \leqslant \lambda \,\mathrm{ES}^g(X) + (1-\lambda)\,\mathrm{ES}^g(Y)$ for all $X, Y \in \mathcal{X}$ and $\lambda \in [0,1]$.*

(5) consistency with $\geqslant_{\mathrm{SSD}}$*: $\mathrm{ES}^g(X) \leqslant \mathrm{ES}^g(Y)$ for all $X, Y \in \mathcal{X}$ such that $X \geqslant_{\mathrm{SSD}} Y$.*

(6) normalization*: $\mathrm{ES}^g(0) = 0$ if and only if $g(0) = 0$.*

It is well known that, in addition to convexity, ES satisfies positive homogeneity. This qualifies it as a coherent risk measure in the sense of Artzner et al. (1999). In the next proposition we show that ES^g satisfies positive homogeneity, i.e.

$$\mathrm{ES}^g(\lambda X) = \lambda \,\mathrm{ES}^g(X) \quad \text{for all } X \in \mathcal{X} \text{ and } \lambda \in (0,\infty),$$

only in the case where it coincides with some ES. In other words, with the exception of ES, the class of adjusted ES's consists of monetary risk measures that are convex but not coherent.

Proposition 2.5. *For every risk profile $g \in \mathcal{G}$ the following statements are equivalent:*

(a) ES^g is positively homogeneous.

(b) $g(0) = 0$ and $g(p) \in (0,\infty)$ for at most one $p \in (0,1]$.

(c) $\mathrm{ES}^g = \mathrm{ES}_p$ where $p = \sup\{q \in [0,1] \mid g(q) = 0\}$.

Proof. "(a)$\Rightarrow$(b)": Since ES^g is positively homogeneous we have

$$\lambda g(0) = -\lambda \,\mathrm{ES}^g(0) = -\,\mathrm{ES}^g(\lambda 0) = -\,\mathrm{ES}^g(0) = g(0)$$

for every $\lambda \in (0,\infty)$. As $g(0) < \infty$ by our assumptions on the class $\mathcal{G}$, we must have $g(0) = 0$. Now, assume by way of contradiction that $0 < g(p_1) \leqslant g(p_2) < \infty$ for some $0 < p_1 < p_2 \leqslant 1$. Take now $q \in (p_1, p_2)$ and $b \in (0, g(p_1))$ and set

$$a = \min\left\{-\frac{(1-q)b}{p-p_1}, \inf_{p\in[0,p_1)} \frac{(1-p)g(p) - b(1-q)}{q-p}\right\}.$$

Note that $a \in (-\infty, 0)$. Since the underlying probability space is assumed to be atomless, we can always find a random variable $X \in \mathcal{X}$ such that

$$F_X(x) = \begin{cases} 0 & \text{if } x \in (-\infty, a), \\ q & \text{if } x \in [a, b), \\ 1 & \text{if } x \in [b, \infty). \end{cases}$$

Note that, for every $p \in [0, p_1)$, the definition of a implies

$$\frac{(1-p)g(p) - b(1-q)}{q-p} \geqslant a.$$

Moreover, for every $p \in [p_1, q)$, the choice of b implies

$$\frac{(1-p)g(p) - b(1-q)}{q-p} \geqslant \frac{(1-p)g(p_1) - b(1-q)}{q-p} \geqslant \frac{(1-p)b - b(1-q)}{q-p} = b \geqslant a.$$

As a result, for every $p \in [0, q)$ we obtain

$$\mathrm{ES}_p(X) = \frac{a(q-p) + b(1-q)}{1-p} \leqslant g(p).$$

Similarly, for every $p \in [q, 1]$ we easily see that

$$\mathrm{ES}_p(X) = b < g(p_1) \leqslant g(q) \leqslant g(p).$$

This yields $\mathrm{ES}^g(X) \leqslant 0$. However, taking $\lambda \in (0, \infty)$ large enough delivers

$$\mathrm{ES}^g(\lambda X) = \sup_{p \in [0,1]} \{\lambda\, \mathrm{ES}_p(X) - g(p)\} \geqslant \lambda\, \mathrm{ES}_q(X) - g(q) = \lambda b - g(q) > 0$$

in contrast to positive homogeneity. As a consequence, we must have $p_1 = p_2$ and thus (b) holds.

"(b)⇒(c)": Set $q = \sup\{p \in [0,1] \mid g(p) = 0\}$. Note that $q \in [0,1]$. Clearly, we have $g(p) = 0$ for every $p \in [0, q)$ and $g(p) = \infty$ for every $p \in (q, 1]$ by assumption. Then, for every $X \in \mathcal{X}$ we get

$$\mathrm{ES}^g(X) = \sup_{p \in [0,q]} \{\mathrm{ES}_p(X) - g(p)\} = \sup_{p \in [0,q]} \mathrm{ES}_p(X) = \mathrm{ES}_q(X)$$

by the continuity of $\mathrm{ES}_p(X)$ in p.

"(c)⇒(a)": The implication is clear. □

Even if positive homogeneity is generally not fulfilled, the following weaker scaling property is satisfied by every adjusted ES because of convexity.

Proposition 2.6. *Consider a risk profile $g \in \mathcal{G}$ with $g(0) = 0$. Then, for every $X \in \mathcal{X}$ the following statement hold:*

(i) $\mathrm{ES}^g(\lambda X) \geqslant \lambda\, \mathrm{ES}^g(X)$ *for every* $\lambda \in [1, \infty)$.

(ii) $\mathrm{ES}^g(\lambda X) \leqslant \lambda\, \mathrm{ES}^g(X)$ *for every* $\lambda \in [0, 1]$.

Let $p \in (0,1)$. We say that a functional $\rho : \mathcal{X} \to (-\infty, \infty]$ satisfies the p-tail property if $\rho(X) = \rho(Y)$ for all random variables $X, Y \in \mathcal{X}$ such that $\mathrm{VaR}_q(X) = \mathrm{VaR}_q(Y)$ for every $q \in [p, 1)$. In words, the p-tail property says that ρ is soely determined by the tail distribution of random losses, and it does not distinguish between two random losses having the same (left) quantiles beyond the level p. This property was introduced by Liu and Wang (2020) to formalize the consideration of tail risk of BCBS (2012). For instance, ES_p satisfies the p-tail property, and the risk measure ES^g in (1) satisfies the 0.99-tail property. Our next result shows characterize this property with a simple condition on g.

Proposition 2.7. *Consider a risk profile* $g \in \mathcal{G}$. *For every* $p \in (0,1)$ *the following statements are equivalent:*

(i) ES^g *satisfies the* p*-tail property.*

(ii) g *is constant on* $(0, p)$.

Proof. To show that (i) implies (ii), assume that g is not constant on $(0,p)$ so that $g(0) < g(r)$ for some $r \in (0,p)$. Without loss of generality we can assume that $g(0) < g(r - \varepsilon)$ for a suitable $\varepsilon > 0$. Now, consider two random variables $X = 1$ and $Y = 1_E$ for some event $E \in \mathcal{F}$ satisfying $\mathbb{P}(E) = r$. Note that $\mathrm{VaR}_q(X) = \mathrm{VaR}_q(Y) = 1$ for every $q \in [p, 1)$. A simple calculation shows that

$$\mathrm{ES}^g(X) = 1 - g(0), \qquad \mathrm{ES}^g(Y) = \max\left\{ \sup_{q \in [0,r)} \left\{ \frac{1-r}{1-q} - g(q) \right\}, 1 - g(r) \right\}.$$

Note that $1 - g(r) < 1 - g(0)$. Moreover, we have

$$\sup_{q \in [0, r-\varepsilon]} \left\{ \frac{1-r}{1-q} - g(q) \right\} \leqslant \frac{1-r}{1-(r-\varepsilon)} - g(0) < 1 - g(0),$$

$$\sup_{q \in (r-\varepsilon, r)} \left\{ \frac{1-r}{1-q} - g(q) \right\} \leqslant 1 - g(r-\varepsilon) < 1 - g(0).$$

This shows that $\mathrm{ES}^g(X) \neq \mathrm{ES}^g(Y)$ and, hence, ES^g fails to satisfy the p-tail property.

To show that (ii) implies (i), assume that g is constant on $(0,p)$. Then, we readily see that

$$\begin{aligned}\mathrm{ES}^g(X) &= \max\left\{ \sup_{q \in [0,p)} \{\mathrm{ES}_q(X) - g(0)\}, \sup_{q \in [p,1]} \{\mathrm{ES}_q(X) - g(q)\} \right\} \\ &= \max\left\{ \mathrm{ES}_p(X) - g(0), \sup_{q \in [p,1]} \{\mathrm{ES}_q(X) - g(q)\} \right\}\end{aligned}$$

for every $X \in \mathcal{X}$. This shows that ES^g satisfies the p-tail property. □

As a next step, we focus on infimal convolutions of adjusted ES's. Infimal convolutions arise naturally in the study of optimal risk sharing and capital allocation problems and have been extensively investigated in the risk measure literature; see, e.g., Barrieu and El Karoui (2005), Burgert and Rüschendorf (2008), Filipović and Svindland (2008) for results in the convex world and Embrechts et al. (2018) for results outside the convex world.

Definition 2.8. Let $n \in \mathbb{N}$ and consider the functionals $\rho_1, \dots, \rho_n : \mathcal{X} \to (-\infty, \infty]$. For every $X \in \mathcal{X}$ we set

$$\mathcal{S}^n(X) := \left\{ (X_1, \dots, X_n) \in \mathcal{X}^n \,\middle|\, \sum_{i=1}^n X_i = X \right\}.$$

The map $\mathop{\square}\limits_{i=1}^{n} \rho_i : \mathcal{X} \to [-\infty, \infty]$ defined by

$$\mathop{\square}\limits_{i=1}^{n} \rho_i(X) := \inf \left\{ \sum_{i=1}^n \rho_i(X_i) \,\middle|\, (X_1, \dots, X_n) \in \mathcal{S}^n(X) \right\},$$

is called the *inf-convolution* of $\{\rho_1, \dots, \rho_n\}$. For $n = 2$ we simply write $\rho_1 \square \rho_2$.

Remark 2.9. It is not difficult to see that, if $\rho_1, \dots, \rho_n : \mathcal{X} \to (-\infty, \infty]$ are monetary risk measures, then for every $X \in \mathcal{X}$ we have

$$\mathop{\square}\limits_{i=1}^{n} \rho_i(X) = \inf\{m \in \mathbb{R}\,;\ X - m \in \mathcal{A}_1 + \cdots + \mathcal{A}_n\}$$

where $\mathcal{A}_i = \{X \in \mathcal{X} \mid \rho_i(X) \leqslant 0\}$ is the acceptance sets induced by ρ_i for $i \in \{1, \dots, n\}$. This shows that the infimal convolution of monetary risk measures is also a monetary risk measure.

We start by establishing a general inequality for inf-convolutions. More precisely, we show that any inf-convolution of adjusted ES's can be controlled from below by a suitable adjusted ES.

Lemma 2.10. *Let $n \in \mathbb{N}$ and consider the risk profiles $g_1, \dots, g_n \in \mathcal{G}$. For every $X \in \mathcal{X}$ we have*

$$\mathop{\square}\limits_{i=1}^{n} \mathrm{ES}^{g_i}(X) \geqslant \mathrm{ES}^{\sum_{i=1}^n g_i}(X).$$

Proof. Clearly, if suffices to focus on the case $n = 2$. For all $Y \in \mathcal{X}$ and $p \in [0, 1]$ we have

$$\mathrm{ES}^{g_1}(Y) + \mathrm{ES}^{g_2}(X - Y) \geqslant \mathrm{ES}_p(Y) - g_1(p) + \mathrm{ES}_p(X - Y) - g_2(p) \geqslant \mathrm{ES}_p(X) - (g_1 + g_2)(p)$$

by the subadditivity of ES. Taking the supremum over p and then the infimum over Y delivers the desired inequality. □

The preceding general inequality can be used to derive a formula for the inf-convolution of an adjusted ES with itself.

Proposition 2.11. *Let $n \in \mathbb{N}$ and consider a risk profile $g \in \mathcal{G}$. For every $X \in \mathcal{X}$ we have*

$$\mathop{\square}_{i=1}^{n} \mathrm{ES}^g(X) = \mathrm{ES}^{ng}(X).$$

Proof. The inequality "$\geqslant$" follows directly from Lemma 2.10. To show the inequality "$\leqslant$", take an arbitrary $X \in \mathcal{X}$ and observe that

$$\mathrm{ES}^g\left(\frac{1}{n}X\right) = \frac{1}{n}\sup_{p\in[0,1]}\{\mathrm{ES}_p(X) - ng(p)\} = \frac{1}{n}\mathrm{ES}^{ng}(X).$$

As a result, we infer that

$$\mathop{\square}_{i=1}^{n} \mathrm{ES}^g(X) \leqslant \sum_{i=1}^{n} \mathrm{ES}^g\left(\frac{1}{n}X\right) = \mathrm{ES}^{ng}(X).$$

This yields the desired inequality and concludes the proof. □

The preceding formula allows us to infer that adjusted ES's exhibit limited regulatory arbitrage in the sense of Wang (2016). As a preliminary step, we recall the notion of regulatory arbitrage in the next definition. (The original definition was in the context of bounded positions and finite risk measures). Recall our convention $\infty - \infty = -\infty$.

Definition 2.12. Consider a functional $\rho : \mathcal{X} \to (-\infty, \infty]$ and for every $X \in \mathcal{X}$ set

$$\underline{\rho}(X) := \inf_{n\in\mathbb{N}} \mathop{\square}_{i=1}^{n} \rho(X).$$

We say that:

(1) ρ is *free of regulatory arbitrage* if $\rho(X) - \underline{\rho}(X) = 0$ for every $X \in \mathcal{X}$.

(2) ρ has *limited regulatory arbitrage* if $\rho(X) - \underline{\rho}(X) < \infty$ for every $X \in \mathcal{X}$.

(3) ρ has *infinite regulatory arbitrage* if $\rho(X) - \underline{\rho}(X) = \infty$ for every $X \in \mathcal{X}$.

It is clear that a risk measure will always exhibit regulatory arbitrage unless it is subadditive. If subadditivity is violated, then the risk measure may incentivize the (internal) reallocation of risk with the only purpose of reaching a lower level of capital requirements. It follows from Proposition 2.5 that adjusted ES's are not subadditive in general and, hence, they will allow for regulatory arbitrage. The next proposition shows that that happens only in a limited form. (The statement for bounded positions follows from Corollary 3.5 in Wang (2016). Note that, in a bounded setting, ES^g is always finite).

Proposition 2.13. *Consider a risk profile $g \in \mathcal{G}$ such that $g(0) = 0$. The following statements hold:*

(i) $\mathrm{ES}^g(X) - \underline{\mathrm{ES}^g}(X) < \infty$ *for every* $X \in \mathcal{X}$ *with* $\mathrm{ES}^g(X) < \infty$.

(ii) $\mathrm{ES}^g(X) - \underline{\mathrm{ES}^g}(X) = \infty$ *for every* $X \in \mathcal{X}$ *with* $\mathrm{ES}^g(X) = \infty$.

Proof. Let $X \in \mathcal{X}$. It follows from Proposition 2.11 that

$$\underline{\rho}(X) = \inf_{n\in\mathbb{N}} \mathrm{ES}^{ng}(X) \geqslant \mathrm{ES}_0(X) = \mathbb{E}[X] > -\infty,$$

where we used that $g(0) = 0$. This delivers the desired statements. □

The ability to express a risk measure in dual terms as a supremum of affine functionals proves a very useful tool in many applications, notably optimization problems; see the general discussion in Rockafellar (1974) and the results on risk measures in Föllmer and Schied (2016). We conclude this section by establishing a dual representation of adjusted ES's. In what follows we denote by $\mathcal{P}$ the set of probability measures on $(\Omega, \mathcal{F})$ and we use the standard notation for Radon-Nikodym derivatives.

Proposition 2.14. *Consider a risk profile* $g \in \mathcal{G}$. *For every* $X \in \mathcal{X}$ *we have*

$$\mathrm{ES}^g(X) = \sup_{\mathbb{Q}\in\mathcal{P}^\infty_\mathbb{P}} \left\{ \mathbb{E}_\mathbb{Q}[X] - g\left(1 - \left\|\frac{d\mathbb{Q}}{d\mathbb{P}}\right\|_\infty^{-1}\right)\right\},$$

where $\mathcal{P}^\infty_\mathbb{P} = \{\mathbb{Q} \in \mathcal{P} \mid \mathbb{Q} \ll \mathbb{P},\ d\mathbb{Q}/d\mathbb{P} \in L^\infty\}$.

Proof. For notational convenience, for every $\mathbb{Q} \in \mathcal{P}^\infty_\mathbb{P}$ set

$$D(\mathbb{Q}) = \left\{ p \in [0,1] \,\middle|\, \frac{d\mathbb{Q}}{d\mathbb{P}} \leqslant \frac{1}{1-p} \right\} = \left[1 - \left\|\frac{d\mathbb{Q}}{d\mathbb{P}}\right\|_\infty^{-1}, 1\right].$$

Take $X \in \mathcal{X}$. The well-known dual representation of ES states that

$$\mathrm{ES}_p(X) = \sup\left\{ \mathbb{E}_\mathbb{Q}[X] \,\middle|\, \mathbb{Q} \in \mathcal{P}^\infty_\mathbb{P},\ \tfrac{d\mathbb{Q}}{d\mathbb{P}} \leqslant \tfrac{1}{1-p} \right\}$$

for every $p \in [0,1]$; see, e.g., Föllmer and Schied (2016). Then, it follows that

$$\begin{aligned}
\mathrm{ES}^g(X) &= \sup_{p\in[0,1]} \left\{ \sup_{\mathbb{Q}\in\mathcal{P}^\infty_\mathbb{P},\ p\in D(\mathbb{Q})} \{\mathbb{E}_\mathbb{Q}[X] - g(p)\} \right\} \\
&= \sup_{\mathbb{Q}\in\mathcal{P}^\infty_\mathbb{P}} \left\{ \sup_{p\in D(\mathbb{Q})} \{\mathbb{E}_\mathbb{Q}[X] - g(p)\} \right\} \\
&= \sup_{\mathbb{Q}\in\mathcal{P}^\infty_\mathbb{P}} \left\{ \mathbb{E}_\mathbb{Q}[X] - \inf_{p\in D(\mathbb{Q})} g(p) \right\}.
\end{aligned}$$

It remains to observe that the above infimum equals $g(1 - \|d\mathbb{Q}/d\mathbb{P}\|_\infty^{-1})$ by monotonicity of g. □

3 Adjusting ES via benchmark ES profiles

In this section we focus on a special class of adjusted ES's for which the target risk profiles are expressed in terms of the ES profile of a reference random loss. As shown below, these special adjusted ES's are intimately linked with monetary risk measures induced by second-order stochastic dominance.

Definition 3.1. Consider a functional $\rho : \mathcal{X} \to (-\infty, \infty]$.

(1) ρ is called a *benchmark-adjusted ES* if there exists $Z \in \mathcal{X}$ such that for every $X \in \mathcal{X}$

$$\rho(X) = \sup_{p \in [0,1]} \{\mathrm{ES}_p(X) - \mathrm{ES}_p(Z)\}.$$

(2) ρ is called an *SSD-based risk measure* if there exists $Z \in \mathcal{X}$ such that for every $X \in \mathcal{X}$

$$\rho(X) = \inf\{m \in \mathbb{R} \mid X - m \geqslant_{\mathrm{SSD}} Z\}.$$

It is clear that benchmark-adjusted ES's are special instances of adjusted ES's for which the target risk profile is defined in terms of the ES profile of a benchmark random loss. The distribution of this random loss may correspond, for example, to the (stressed) historical loss distribution of the underlying position or to a target (risk-class specific) loss distribution. It is also clear that SSD-based risk measures are nothing but monetary risk measures associated with acceptance sets defined through second-order stochastic dominance.

The classical characterization of second-order stochastic dominance in terms of ES can be used to show that benchmark-adjusted ES's coincide with SSD-based risk measures. In addition, we provide a simple characterization of this class of risk measures.

Theorem 3.2. *For a monetary risk measure $\rho : \mathcal{X} \to (-\infty, \infty]$ the following statements are equivalent:*

(i) ρ is a benchmark-adjusted ES.

(ii) ρ is an SSD-based risk measure.

(iii) ρ is consistent with $\geqslant_{\mathrm{SSD}}$ and the set $\{X \in \mathcal{X} \mid \rho(X) \leqslant 0\}$ has an $\geqslant_{\mathrm{SSD}}$-minimum element.

Proof. Recall that for all $X \in \mathcal{X}$ and $Z \in \mathcal{X}$ we have $X \geqslant_{\mathrm{SSD}} Z$ if and only if $\mathrm{ES}_p(X) \leqslant \mathrm{ES}_p(Z)$ for every $p \in [0,1]$; see, e.g., Theorem 4.A.3 in Shaked and Shanthikumar (2007). For convenience, set $\mathcal{A} = \{X \in \mathcal{X} \mid \rho(X) \leqslant 0\}$. To show that (i) implies (ii), assume that ρ is a benchmark-adjusted ES with respect to $Z \in \mathcal{X}$. Then, for every $X \in \mathcal{X}$

$$\begin{aligned}\rho(X) &= \inf\{m \in \mathbb{R} \mid X - m \in \mathcal{A}\} \\ &= \inf\{m \in \mathbb{R} \mid \mathrm{ES}_p(X) - m \leqslant \mathrm{ES}_p(Z),\ \forall p \in [0,1]\} \\ &= \inf\{m \in \mathbb{R} \mid X - m \geqslant_{\mathrm{SSD}} Z\}.\end{aligned}$$

To show that (ii) implies (i), assume that ρ is SSD-based with respect to $Z \in \mathcal{X}$. Then, we have

$$\begin{aligned}\rho(X) &= \inf\{m \in \mathbb{R} \mid X - m \geqslant_{\mathrm{SSD}} Z\} \\ &= \inf\{m \in \mathbb{R} \mid \mathrm{ES}_p(X) - m \leqslant \mathrm{ES}_p(Z),\ \forall p \in [0,1]\} \\ &= \sup_{p \in [0,1]} \{\mathrm{ES}_p(X) - \mathrm{ES}_p(Z)\}.\end{aligned}$$

It is clear that (iii) implies (ii). Finally, to show that (ii) implies (iii), assume that ρ is an SSD-based risk measure with respect to $Z \in \mathcal{X}$. It is clear that $Z \in \mathcal{A}$. Now, take an arbitrary $X \in \mathcal{A}$. We find a sequence $(m_n) \subset \mathbb{R}$ such that $m_n \downarrow \rho(X)$ and $X - m_n \geqslant_{\mathrm{SSD}} Z$ for every $n \in \mathbb{N}$. This implies that $X - \rho(X) \geqslant_{\mathrm{SSD}} Z$. Since $\rho(X) \leqslant 0$, we infer that $X \geqslant_{\mathrm{SSD}} Z$ as well. This shows that $\mathcal{A}$ has an SSD-minimum element. To establish that ρ is consistent with $\geqslant_{\mathrm{SSD}}$, take arbitrary $X, Y \in \mathcal{X}$ satisfying $X \geqslant_{\mathrm{SSD}} Y$. For every $m \in \mathbb{R}$ such that $Y - m \geqslant_{\mathrm{SSD}} Z$ we clearly have that $X - m \geqslant_{\mathrm{SSD}} Y - m \geqslant_{\mathrm{SSD}} Z$. This implies that $\rho(X) \leqslant \rho(Y)$ and concludes the proof of the desired implication. □

Remark 3.3. (i) Let $\mathcal{L}$ be the family of all (nonconstant) convex and increasing functions $\ell : \mathbb{R} \to \mathbb{R}$. The monetary risk measure associated to $\ell \in \mathcal{L}$ is defined for a given $\alpha \in \mathbb{R}$ by

$$\rho_{\ell,\alpha}(X) := \inf\{m \in \mathbb{R} \mid \mathbb{E}[\ell(X - m)] \leqslant \alpha\}, \quad X \in \mathcal{X}.$$

Consider the risk profile $g(p) = \mathrm{ES}_p(Z)$ for every $p \in [0,1]$, where Z is a given $\mathbb{P}$-essentially bounded random variable. Then, it is not difficult to verify that

$$\mathrm{ES}^g(X) = \sup_{\ell \in \mathcal{L}} \rho_{\ell, \mathbb{E}[\ell(Z)]}(X)$$

for every $X \in \mathcal{X}$. In particular, ES^g is more conservative than any shortfall risk measure with reference level $\mathbb{E}[\ell(Z)]$.

(ii) If second-order stochastic dominance is replaced by first-order stochastic dominance in the above theorem, then one arrives at a characterization of LVaR in (2) with α being a distribution function.

We are interested in characterizing when the acceptable risk profile g of an adjusted ES can be expressed in terms of an ES profile. To this effect, it is convenient to introduce the following additional class of risk measures, which will be shown to contain all benchmark-adjusted ES's.

Definition 3.4. A functional $\rho : \mathcal{X} \to (-\infty, \infty]$ is called a *quantile-adjusted ES* if there exists $Z \in L^0$ such that for every $X \in \mathcal{X}$

$$\rho(X) = \sup_{p\in[0,1]} \{\mathrm{ES}_p(X) - \mathrm{VaR}_p(Z)\}.$$

To establish our desired characterization, for a risk profile $g \in \mathcal{G}$ we define $h_g : [0,1] \to (-\infty, \infty]$ by

$$h_g(p) := (1-p)g(p).$$

Here, we set $0 \cdot \infty = 0$ so that $h_g(1) = 0$. Moreover, we introduce the following sets:

$$\mathcal{G}_{\mathrm{VaR}} := \{g \in \mathcal{G} \mid g \text{ is finite on } [0,1), \text{ left-continuous on } [0,1], \text{ and right-continuous at } 0\},$$

$$\mathcal{G}_{\mathrm{ES}} := \{g \in \mathcal{G}_{\mathrm{VaR}} \mid h_g \text{ is concave on } (0,1) \text{ and left-continuous at } 1\}.$$

Lemma 3.5. *For every risk profile $g \in \mathcal{G}$ the following statements hold:*

(i) $g \in \mathcal{G}_{\mathrm{VaR}}$ if and only if there exists a random variable $Z \in L^0$ that is bounded from below and satisfies $g(p) = \mathrm{VaR}_p(Z)$ for every $p \in [0,1]$.

(ii) $g \in \mathcal{G}_{\mathrm{ES}}$ if and only if there exists a random variable $Z \in \mathcal{X}$ such that $g(p) = \mathrm{ES}_p(Z)$ for every $p \in [0,1]$.

Proof. (i) The "if" part is clear. For the "only if" part, let U be a uniform random variable on $[0,1]$ and set $Z = g(U)$. Then, it is well known that $\mathrm{VaR}_p(Z) = g(p)$ for every $p \in [0,1]$. Moreover, since $g(0) > -\infty$, we see that Z is bounded from below.

(ii) The "if" part is straightforward. For the "only if" part, let U be a uniform random variable on $[0,1]$. We denote by h'_g the left derivative of h_g. Then, for every $p \in [0,1)$ we have

$$\mathrm{ES}_p(-h'_g(U)) = -\frac{1}{1-p}\int_p^1 h'_g(u)\mathrm{d}u = -\frac{h_g(1) - h_g(p)}{1-p} = g(p).$$

This shows that, by taking $Z = -h'_g(U)$, we have $g(p) = \mathrm{ES}_p(Z)$ for every $p \in [0,1)$. The left continuity of g and $\mathrm{ES}_\cdot(Z)$ at 1 gives the same equality for $p = 1$. □

As a direct consequence of the previous lemma we derive a characterization of quantile- and benchmark-adjusted ES's in terms of the underlying risk profile.

Theorem 3.6. *For every risk profile $g \in \mathcal{G}$ the following statements hold:*

(i) There exists $Z \in L^0$ that is bounded from below and such that ES^g is a quantile-adjusted ES with respect to Z if and only if $g \in \mathcal{G}_{\mathrm{VaR}}$.

(ii) There exists $Z \in \mathcal{X}$ such that ES^g is an benchmark-adjusted ES with respect to Z if and only if $g \in \mathcal{G}_{\mathrm{ES}}$.

Since we clearly have $\mathcal{G}_{\mathrm{ES}} \subset \mathcal{G}_{\mathrm{VaR}}$, it follows from the above results that every benchmark-adjusted ES is also a quantile-adjusted ES. In particular, this implies that, for every random variable $Z \in \mathcal{X}$, we can always find a random variable $W \in L^0$ such that $\mathrm{VaR}_p(W) = \mathrm{ES}_p(Z)$ for every $p \in [0, 1]$. In words, every ES profile can be reproduced by a suitable VaR profile. As pointed out by the next proposition, the converse result is, in general, not true. In addition, we also show that an adjusted ES need not be a quantile-adjusted ES.

Proposition 3.7. *(i) There exists $g \in \mathcal{G}$ such that $\mathrm{ES}^g \neq \mathrm{ES}^h$ for every $h \in \mathcal{G}_{\mathrm{VaR}}$.*

(ii) There exists $g \in \mathcal{G}_{\mathrm{VaR}}$ such that $\mathrm{ES}^g \neq \mathrm{ES}^h$ for every $h \in \mathcal{G}_{\mathrm{ES}}$.

Proof. The second assertion follows immediately from Theorem 3.6 and the fact that the inclusion $\mathcal{G}_{\mathrm{ES}} \subset \mathcal{G}_{\mathrm{VaR}}$ is strict. To establish the first assertion, fix $q \in (0, 1)$ and define $g \in \mathcal{G}$ by setting

$$g(p) = \begin{cases} 0 & \text{if } p \in [0, q], \\ \infty & \text{if } p \in (q, 1]. \end{cases}$$

It follows that

$$\mathrm{ES}^g(X) = \sup_{p \in [0,q]} \{\mathrm{ES}_p(X)\} = \mathrm{ES}_q(X)$$

for every $X \in \mathcal{X}$. We claim that ES^g is not a quantile-adjusted ES. To the contrary, suppose that there exists a random variable $Z \in L^0$ that is bounded from below and satisfies

$$\mathrm{ES}_q(X) = \mathrm{ES}^g(X) = \sup_{p \in [0,1]} \{\mathrm{ES}_p(X) - \mathrm{VaR}_p(Z)\}$$

for every $X \in \mathcal{X}$. Take $r \in (q, 1)$ and $X \in \mathcal{X}$ such that $\mathrm{ES}_r(X) > \mathrm{ES}_q(X)$. Then, for each $\lambda > 0$

$$\begin{aligned} \mathrm{ES}_q(X) &= \frac{1}{\lambda} \mathrm{ES}_q(\lambda X) = \frac{1}{\lambda} \sup_{p \in [0,1]} \{\mathrm{ES}_p(\lambda X) - \mathrm{VaR}_p(Z)\} \\ &\geqslant \frac{1}{\lambda}(\mathrm{ES}_r(\lambda X) - \mathrm{VaR}_r(Z)) = \mathrm{ES}_r(X) - \frac{1}{\lambda} \mathrm{VaR}_r(Z). \end{aligned}$$

By sending $\lambda \to \infty$, we obtain $\mathrm{ES}_q(X) \geqslant \mathrm{ES}_r(X)$, which contradicts our assumption on X. □

Note that ES is always finite on our domain. Here, we are interested in discussing the finiteness of adjusted ES's associated with risk profiles in the class $\mathcal{G}_{\mathrm{VaR}}$ and $\mathcal{G}_{\mathrm{ES}}$. We show that finiteness on the whole reference space $\mathcal{X}$ can never hold in the presence of a risk profile in $\mathcal{G}_{\mathrm{ES}}$ while it can hold if we take a risk profile in $\mathcal{G}_{\mathrm{VaR}}$.

Proposition 3.8. *Consider a risk profile $g \in \mathcal{G}$. If $g \in \mathcal{G}_{\mathrm{VaR}}$, then ES^g can be finite on $\mathcal{X}$. If $g \in \mathcal{G}_{\mathrm{ES}}$, then ES^g cannot be finite on $\mathcal{X}$.*

Proof. To show the first part of the assertion, set $g(p) = \frac{1}{1-p}$ for every $p \in [0,1]$ (with the convention $\frac{1}{0} = \infty$). Note that $g \in \mathcal{G}_{\mathrm{VaR}}$. Fix $X \in \mathcal{X}$ and note that there exists $q \in (0,1)$ such that

$$\sup_{p\in[q,1]} \int_p^1 \mathrm{VaR}_r(X)\mathrm{d}r < 1.$$

It follows that

$$\sup_{p\in[q,1]} \left\{\mathrm{ES}_p(X) - \frac{1}{1-p}\right\} = \sup_{p\in[q,1]} \left\{\frac{1}{1-p}\left(\int_p^1 \mathrm{VaR}_r(X)\mathrm{d}r - 1\right)\right\} \leqslant 0.$$

Therefore,

$$\mathrm{ES}^g(X) \leqslant \max\left\{\sup_{p\in[0,q]} \left\{\mathrm{ES}_p(X) - \frac{1}{1-p}\right\}, 0\right\} \leqslant \max\{\mathrm{ES}_q(X), 0\} < \infty.$$

This shows that ES^g is finite on the entire $\mathcal{X}$. To establish the second part of the assertion, take $Z \in \mathcal{X}$ and set $g(p) = \mathrm{ES}_p(Z)$ for every $p \in [0,1]$. Note that $g \in \mathcal{G}_{\mathrm{ES}}$ by Lemma 3.5. If Z is bounded from above, then take $X \in \mathcal{X}$ that is unbounded from above. In this case, it follows that

$$\mathrm{ES}^g(X) \geqslant \mathrm{ES}_1(X) - \mathrm{ES}_1(Z) = \infty.$$

If Z is unbounded from above, then take $X = 2Z \in \mathcal{X}$. In this case, we have

$$\mathrm{ES}^g(X) \geqslant \mathrm{ES}_1(2Z) - \mathrm{ES}_1(Z) = \mathrm{ES}_1(Z) = \infty.$$

Hence, we see that ES^g is never finite on $\mathcal{X}$. □

The next result improves Lemma 2.10 by showing that the inf-convolution of benchmark-adjusted ES's can still be expressed as an adjusted ES.

Proposition 3.9. *Let $n \in \mathbb{N}$ and consider the risk profiles $g_1, \ldots, g_n \in \mathcal{G}_{\mathrm{ES}}$. For every $X \in \mathcal{X}$*

$$\mathop{\square}_{i=1}^{n} \mathrm{ES}^{g_i}(X) = \mathrm{ES}^{\sum_{i=1}^n g_i}(X).$$

Proof. The inequality "$\geqslant$" follows from Lemma 2.10. To show the inequality "$\leqslant$", note that there exist $Z_1, \dots, Z_n \in \mathcal{X}$ such that $\mathcal{A}_{g_i} = \{X \in \mathcal{X} \mid X \geqslant_{\mathrm{SSD}} Z_i\}$ by Theorem 3.6. We prove that

$$\mathcal{A} := \{X \in \mathcal{X}\ ;\ \mathrm{ES}^{\sum_{i=1}^n g_i}(X) \leqslant 0\} \subset \sum_{i=1}^n \mathcal{A}_{g_i}$$

which, together with Remark 2.9, yields the desired inequality. Let U be a uniform random variable and, for any $X \in \mathcal{X}$, denote by F_X^{-1} the (left) quantile function of X. Take $i \in \{1, \dots, n\}$ and note that $F_{Z_i}^{-1}(U) \sim Z_i$. It follows from the law invariance of ES that $\mathrm{ES}_p(F_{Z_i}^{-1}(U)) = \mathrm{ES}_p(Z_i)$ for every $p \in [0,1]$, so that $F_{Z_i}^{-1}(U) \in \mathcal{A}_{g_i}$. Since the random variables $F_{Z_i}^{-1}(U)$'s are comonotonic, we also have $\sum_{i=1}^n \mathrm{ES}_p(Z_i) = \sum_{i=1}^n \mathrm{ES}_p(F_{Z_i}^{-1}(U)) = \mathrm{ES}_p(Z)$ with $Z = \sum_{i=1}^n F_{Z_i}^{-1}(U)$. We deduce that each $X \in \mathcal{A}$ satisfies $\mathrm{ES}_p(X) \leqslant \mathrm{ES}_p(Z)$ for every $p \in [0,1]$, which is equivalent to $X \geqslant_{\mathrm{SSD}} Z$. Note that $Z \in \sum_{i=1}^n \mathcal{A}_{g_i}$ so that $\mathop{\square}_{i=1}^{n} \mathrm{ES}^{g_i}(Z) \leqslant 0$. Since the inf-convolution is consistent with $\geqslant_{\mathrm{SSD}}$, as shown in Theorem 4.1 by Mao and Wang (2020), we have $\mathop{\square}_{i=1}^{n} \mathrm{ES}^{g_i}(X) \leqslant \mathop{\square}_{i=1}^{n} \mathrm{ES}^{g_i}(Z) \leqslant 0$, which implies $X \in \sum_{i=1}^n \mathcal{A}_{g_i}$ as desired. □

4 Optimization with benchmark-adjusted ES's

Using the characterization of benchmark-adjusted ES's established in Theorem 3.2, many optimization problems related to benchmark-adjusted ES's or, equivalently, SSD-based risk measures can be solved explicitly. In this section, we focus on risk minimization and utility maximization problems in the context of a multi-period frictionless market that is complete and arbitrage free. The interest rate is set to be zero for simplicity. As is commonly done in the literature, this type of optimization problems, which are naturally expressed in terms of dynamic investment strategies, can be converted into static optimization problems by way of martingale methods. Below we focus directly on their static counterparts. For more details we refer, e.g., to Schied et al. (2009) or Föllmer and Schied (2016). In addition, to ensure that all our problems are well defined, we assume throughout that $\mathcal{X}$ consists of $\mathbb{P}$-bounded random variables.

In the sequel, we denote by $\mathbb{Q}$ the risk-neutral pricing measure (whose existence and uniqueness in our setting are ensured by the Fundamental Theorem of Asset Pricing), by $w \in \mathbb{R}$ a fixed level of initial wealth, by $x \in \mathbb{R}$ a real number representing a constraint, by $u : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ a concave and increasing function that is continuous (at the point where it potentially jumps to $-\infty$) and satisfies $u(-\infty) < x < u(\infty)$, and by $\rho : \mathcal{X} \to (-\infty, \infty]$ a risk functional. We focus on the following five optimization problems:

(i) Risk minimization with a budget constraint:

$$\text{minimize } \rho(X) \text{ over } X \in \mathcal{X} \text{ subject to } \mathbb{E}_{\mathbb{Q}}[w - X] \leqslant x.$$

(ii) Price minimization with controlled risk:

$$\text{minimize } \mathbb{E}_{\mathbb{Q}}[w - X] \text{ over } X \in \mathcal{X} \text{ subject to } \rho(X) \leqslant x.$$

(iii) Risk minimization with a target utility level:

$$\text{minimize } \rho(X) \text{ over } X \in \mathcal{X} \text{ subject to } \mathbb{E}[u(w - X)] = x.$$

(iv) Worst-case utility with a reference risk assessment:

$$\text{minimize } \mathbb{E}[u(w - X)] \text{ over } X \in \mathcal{X} \text{ subject to } \rho(X) = x.$$

(v) Worst-case risk with a reference risk assessment:

$$\text{maximize } \rho'(X) \text{ over } X \in \mathcal{X} \text{ subject to } \rho(X) = x,$$

where ρ' is an SSD-consistent functional that is continuous with respect to the L^∞-norm.

Problem (i) is an optimal investment problem minimizing the risk given a budget constraint. Conversely, problem (ii) aims at minimizing the cost given a controlled risk level. Problem (iii) is about minimizing the risk exposure with a target utility level, similar to the mean-variance problem of Markowitz (1952). The interpretation of problems (iv) and (v) is different from the first three problems: They are not about optimization over risk, but about ambiguity, i.e., in these problems the main concern is model risk. Indeed, the set $\mathcal{X}$ may represent the class of plausible models for the distribution of a certain financial position of interest. In the case of problem (iv), the assumption is that the only available information for X is the risk figure $\rho(X)$, evaluated, e.g., by an expert or another decision maker. In this context, we are interested in determining the worst case utility among all possible models which agree with the evaluation $\rho(X) = x$ (see also Example 5.3 of Wang et al. (2019)). A similar interpretation can be given for problem (v).

Proposition 4.1. *Each of the optimization problems (i)-(v) relative to a benchmark-adjusted ES* $\rho = \mathrm{ES}^g$ *for* $g \in \mathcal{G}_{\mathrm{ES}}$ *admits an optimal solution of the explicit form* $Z + z$ *where* $Z \in \mathcal{X}$ *has the ES profile* g *and* $z \in \mathbb{R}$*. Moreover,* Z *is comonotonic with* $\frac{d\mathbb{Q}}{d\mathbb{P}}$ *in (i)-(ii), and the (binding) constraint uniquely determines* z *in each problem.*

Proof. The result for the optimization problem (i) is a direct consequence of Proposition 5.2 in Mao and Wang (2020). Let Z be comonotonic with $d\mathbb{Q}/d\mathbb{P}$ which has ES profile g; comonotonicity is only relevant in problems (i) and (ii). Note that $\rho(Z)=0$. For any random variable $X\in\mathcal{X}$, we set $Y_X=Z+\rho(X)$. It is clear that $\rho(Y_X)=\rho(X)$ and

$$\mathrm{ES}_p(Y_X)=g(p)+\rho(X)=g(p)+\sup_{q\in[0,1]}\{\mathrm{ES}_q(X)-g(q)\}\geqslant \mathrm{ES}_p(X).$$

Hence, $X\geqslant_{\mathrm{SSD}} Y_X$. This observation will be useful in the analysis below.

(a) We first look at problem (ii). First, since both $X\mapsto\mathbb{E}_{\mathbb{Q}}[X]$ and ρ are translation-invariant, the condition $\rho(X)\leqslant x$ is binding, and problem (ii) is equivalent to maximizing $\mathbb{E}_{\mathbb{Q}}[X]$ over $X\in\mathcal{X}$ such that $\rho(X)=x$. Let $X\in\mathcal{X}$ be any random variable with $\rho(X)=x$ and let X' be identically distributed as X and comonotonic with $d\mathbb{Q}/d\mathbb{P}$. Since $X'\sim X$, by the Hardy–Littlewood inequality (e.g., Remark 3.25 of Rüschendorf (2013)), we have $\mathbb{E}_{\mathbb{Q}}[X]\leqslant\mathbb{E}_{\mathbb{Q}}[X']$. Moreover, for any random variable $Y\in\mathcal{X}$ that is comonotonic with $d\mathbb{Q}/d\mathbb{P}$, we can write (see e.g., (A.8) of Mao and Wang (2020))

$$\mathbb{E}_{\mathbb{Q}}[Y]=\int_0^1\mathrm{ES}_p(Y)\mathrm{d}\mu(p)$$

for some Borel probability measure μ on $[0,1]$. Hence, $X'\geqslant_{\mathrm{SSD}} Y_X$ implies $\mathbb{E}_{\mathbb{Q}}[X']\leqslant\mathbb{E}_{\mathbb{Q}}[Y_X]$, and we obtain

$$\mathbb{E}_{\mathbb{Q}}[X]\leqslant\mathbb{E}_{\mathbb{Q}}[X']\leqslant\mathbb{E}_{\mathbb{Q}}[Y_X].$$

Note also that $\rho(Y_X)=\rho(X)=x$. Hence, for any random variable $X\in\mathcal{X}$, there exists $Z+z$ for some $z\in\mathbb{R}$ which dominates X for problem (ii). Since both the constraint and the objective are continuous in $z\in\mathbb{R}$, an optimizer of the form $Z+z$ exists.

(b) We next look at problem (iii). Let $X\in\mathcal{X}$ be any random variable such that $\mathbb{E}[u(w-X)]=x$. The aforementioned fact $X\geqslant_{\mathrm{SSD}} Y_X$ implies that $\mathbb{E}[u(w-Y)]\leqslant\mathbb{E}[u(w-X)]=x$ since u is a concave utility function. Therefore, there exists $\varepsilon\in[0,\infty)$ such that $\mathbb{E}[u(w-(Y-\varepsilon))]=x$, and we take the largest ε satisfying this equality, which is obviously finite. Let $z=\rho(X)-\varepsilon$. It is then clear that $\mathbb{E}[u(w-(Z+z))]=\mathbb{E}[u(w-X)]=x$ and $\rho(Z+z)=\rho(Y-\varepsilon)=\rho(X)-\varepsilon\leqslant\rho(X)$. Hence, $Z+z$ dominates X as an optimizer for problem (iii). Since both the constraint and the objective are continuous in $z\in\mathbb{R}$, an optimizer of the form $Z+z$ exists.

(c) Problems (iv) and (v) are similar, and they can be shown via similar arguments to the above cases. □

Remark 4.2. (i) It should be clear that the classical ES does not belong to the class of SSD-based risk measures as the associated risk profile is not in $\mathcal{G}_{\mathrm{ES}}$. As a consequence, the results in this section do not directly apply to ES. In particular, although ES is consistent with SSD, its acceptance set does not have a minimum SSD element as required by Proposition 3.2. We refer to Wang and Zitikis (2020) for a different characterization of ES.

(ii) In the context of decision theory and, specifically, portfolio selection, it is sometimes argued that (second order) stochastic dominance is too extreme in the sense that it ranks risks according to the simultaneous preferences of *every* risk-averse agent, thus including utility functions that may lead to counterintuitive outcomes. A typical example is the one proposed by Leshno and Levy (2002). Consider a portfolio that pays one million dollars in 99% of cases and nothing otherwise and another portfolio that pays one dollar with certainty. According to the sign convention adopted in this paper, the corresponding payoffs are given by

$$X = \begin{cases} 0 & \text{with probability } 1\% \\ -10^6 & \text{with probability } 99\% \end{cases} \quad \text{and} \quad Y = -1.$$

Even though X does not dominate Y with respect to SSD, most agents prefer X to Y. Thus, the authors argue for the necessity of relaxing SSD in favor of a more reasonable notion. We point out that our approach yields a novel and reasonable generalization of SSD. First, consider the risk profile defined by $g(p) = \mathrm{ES}_p(Y) = -1$ for every $p \in [0,1]$ and note that X is acceptable under ES^g precisely when $X \geqslant_{\mathrm{SSD}} Y$. Note also that

$$\mathrm{ES}_p(X) \leqslant g(p) \iff p \leqslant \bar{p} := 1 - \frac{10^{-4}}{10^6 - 1} \approx 1 - 10^{-10}.$$

This fact has two implications. On the one hand, it confirms that X does not dominate Y with respect to SSD and highlights that this failure is due to the behavior of X in the far region of its left tail. On the other hand, it suggests that it is enough to consider the new risk profile defined by $h(p) = g(p)$ for $p \leqslant \bar{p}$ and $h(p) = \infty$ otherwise to make X acceptable under ES^h. In other words, moving from g to h is equivalent to moving from SSD to a relaxed form of SSD that enlarges the spectrum of acceptability in portfolio selection problems. However, note that ES^h is not an SSD-based risk measure and, hence, the existence results obtained above do not apply to it. A systematic study of optimization problems under constraints of ES^h type requires further research.

References

Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance*, 26(7):1505–1518.

Acerbi, C. and Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7):1487–1503.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3):203–228.

Baes, M., Koch-Medina, P., and Munari, C. (2020). Existence, uniqueness, and stability of optimal payoffs of eligible assets. *Mathematical Finance*.

Barrieu, P. and El Karoui, N. (2005). Inf-convolution of risk measures and optimal risk transfer. *Finance and Stochastics*, 9(2):269–298.

BCBS (2012). *Consultative Document May 2012. Fundamental review of the trading book.* Basel Committee on Banking Supervision. Basel: Bank for International Settlements.

Biener, C., Eling, M., and Wirfs, J. (2015). Insurability of cyber risk: An empirical analysis. *Geneva Papers on Risk and Insurance - Issues and Practice*, 40:131–158.

Bignozzi, V., Burzoni, M., and Munari, C. (2020). Risk measures based on benchmark loss distributions. *Journal of Risk and Insurance*, 87(2):437–475.

Burgert, C. and Rüschendorf, L. (2008). Allocation of risks and equilibrium in markets with finitely many traders. *Insurance: Mathematics and Economics*, 42(1):177–188.

Cont, R., Deguest, R., and Scandolo, G. (2010). Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance*, 10(6):593–606.

Du, Z. and Escanciano, J. C. (2017). Backtesting expected shortfall: accounting for tail risk. *Management Science*, 63(4):940–958.

Embrechts, P., Liu, H., and Wang, R. (2018). Quantile-based risk sharing. *Operations Research*, 66(4):936–949.

Filipović, D. and Svindland, G. (2008). Optimal capital and risk allocations for law-and cash-invariant convex functions. *Finance and Stochastics*, 12(3):423–439.

Föllmer, H. and Schied, A. (2016). *Stochastic Finance: An Introduction in Discrete Time.* De Gruyter, Berlin, 4th edition.

Frey, R. and McNeil, A. J. (2002). Var and expected shortfall in portfolios of dependent credit risks: conceptual and practical insights. *Journal of banking & finance*, 26(7):1317–1334.

Koch-Medina, P. and Munari, C. (2016). Unexpected shortfalls of expected shortfall: Extreme default profiles and regulatory arbitrage. *Journal of Banking & Finance*, 62:141–151.

Krätschmer, V., Schied, A., and Zähle, H. (2014). Comparative and qualitative robustness for law-invariant risk measures. *Finance and Stochastics*, 18(2):271–295.

Kratz, M., Lok, Y. H., and McNeil, A. J. (2018). Multinomial var backtests: A simple implicit approach to backtesting expected shortfall. *Journal of Banking & Finance*, 88:393–407.

Leshno, M. and Levy, H. (2002). Preferred by "all" and preferred by "most" decision makers: Almost stochastic dominance. *Management Science*, 48(8):1074–1085.

Liu, F. and Wang, R. (2020). A theory for measures of tail risk. *Mathematics of Operations Research, forthcoming.*

Mao, T. and Wang, R. (2020). Risk aversion in regulatory capital principles. *SIAM Journal on Financial Mathematics*, 11(1):169–200.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7:77–91.

Refsdal, A., Solhaug, B., and Stølen, K. (2015). *Cyber-risk Management*, pages 33–47. Springer International Publishing, Cham.

Rockafellar, R. T. (1974). *Conjugate duality and optimization*. SIAM.

Rockafellar, R. T. and Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of banking & finance*, 26(7):1443–1471.

Rüschendorf, L. (2013). Mathematical risk analysis. *Springer Ser. Oper. Res. Financ. Eng. Springer, Heidelberg*.

Schied, A., Föllmer, H., and Weber, S. (2009). Robust preferences and robust portfolio choice. *Handbook of Numerical Analysis*, 15:29–87.

Shaked, M. and Shanthikumar, J. (2007). *Stochastic Orders*. Springer Series in Statistics. Springer New York.

Wang, R. (2016). Regulatory arbitrage of risk measures. *Quantitative Finance*, 16(3):337–347.

Wang, R., Xu, Z. Q., and Zhou, X. Y. (2019). Dual utilities on risk aggregation under dependence uncertainty. *Finance and Stochastics*, 23(4):1025–1048.

Wang, R. and Zitikis, R. (2020). An axiomatic foundation for the expected shortfall. *Management Science*, forthcoming.

Weber, S. (2018). Solvency ii, or how to sweep the downside risk under the carpet. *Insurance: Mathematics and economics*, 82:191–200.

Ziegel, J. F. (2016). Coherence and elicitability. *Mathematical Finance*, 26(4):901–918.